DECISION-TREE AND ENSEMBLE-BASED MORTALITY RISK MODELS FOR HOSPITALIZED PATIENTS WITH COVID-19

Ya. VYKLYUK, S. LEVYTSKA, D. NEVINSKYI, K. HAZDIUK, M. ŠKODA, S. ANDRUSHKO, M. PALII

Abstract. The work is devoted to studying SARS-CoV-2-associated pneumonia and the investigating of the main indicators that lead to the patients' mortality. Using the good-known parameters that are routinely embraced in clinical practice, we obtained new functional dependencies based on an accessible and understandable decision tree and ML ensemble of classifiers models that would allow the physician to determine the prognosis in a few minutes and, accordingly, to understand the need for treatment adjustment, transfer of the patient to the emergency department. The accuracy of the resulting ensemble of models fitted on actual hospital patient data was in the range of 0.88–0.91 for different metrics. Creating a data collection system with further training of classifiers will dynamically increase the forecast's accuracy and automate the doctor's decision-making process.

Keywords: COVID-19, decision-making system, decision tree, ML-ensemble, ensemble of classification models.

BACKGROUND

The pandemic of SARS-CoV-2 infection, started in December 2019 has rapidly spread across the globe and affected all countries in two years. As of November 2021, the number of world-wide cases exceeded 262 million people, more than 5 million people died, including more than 85 thousand deaths in Ukraine [1]. The spread of coronavirus infection in Ukraine began from Chernivtsi and this city held the sad first place by the level of the SARS-CoV-2 morbidity during a year and a half. An emergency situation in medicine has obliged physicians of various specialties to help patients with SARS-CoV-2-associated pneumonia and to study the peculiarities of SARS-CoV-2 infection in their own practical experience.

Despite the huge accumulated clinical and laboratory material, the extraordinary attention of the medical community to the treatment of patients with SARS-CoV-2-associated pneumonia, it is still not clear why the disease became fatal for some patients [2].

Recent years decision-making and expert systems based on artificial intelligence have become widespread in medicine. Classification methods are one of the most urgent and necessary tasks in medicine. Classification shapes medicine and guides its practice. An understanding of classification should be part of the search for a better understanding of the social context and consequences of diagnosis. Classification is the part of human activity that provides the basis for recognizing and studying a disease. This means deciding how to extract significant parts from the vast expanse of nature, stabilizing and structuring disordered things [3], [4]. One of the most popular methods of classification is the diagnosic X-ray.

Different types of convolutional neural networks, or classical classifiers based on image features, are used as a classification model [5–7].

There are also investigations to determine mortality rate of patients depending on medical indicators. In particular, in the paper [8] Lactate dehydrogenase, neutrophils (%), lymphocyte (%), high-sensitivity C-reactive protein, and age (LNLCA), which were determined on hospital admission, were identified as key predictors of death from the multi-tree XGBoost model. The integrated score (LNLCA) was calculated with the corresponding probability of death. COVID-19 patients were divided into three subgroups: low-, middle-, and high-risk groups using LNLCA cutoff values of 10.4 and 12.65. The probability of death in each group is less than 5%, 5-50% and above 50%, respectively. The prognostic model, nomogram, and LNLCA assessment can help identify early high-risk mortality in patients with COVID-19, which will help physicians improve the management of patient stratification.

In the paper [9] the severity and outcome of COVID-19 cases has been associated with the percentage of circulating lymphocytes (LYM%), levels of C-reactive protein (CRP), interleukin-6 (IL-6), procalcitonin (PCT), lactic acid (LA), and viral load (ORF1ab Ct). However, the predictive power of each of these indicators in disease classification and prognosis remains largely unclear.

Similar results in work [10] indicate that the risk period for patients is 12–14 days, after which the probability of patient survival may increase. In addition, it is noted that the probability of death in COVID cases increases with age. It is established that the probability of death is higher in men than in women. SVM with Grid search methods showed the highest accuracy of about 95%, followed by the decision tree algorithm with an accuracy of about 94%.

Retrospective Cohort Study [11] included patients with COVID-19 who were admitted at three designated locations at Wuhan Union Hospital (Wuhan, China). Dynamic hematological and coagulation parameters were investigated with a linear mixed model, and coagulopathy screening with sepsis-induced coagulopathy and International Society of Thrombosis and Hemostasis overt disseminated intravascular coagulation scoring systems was applied.

The authors of paper [12] used the available information on pre-existing health conditions identified for deceased patients positive with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)' in Italy. They estimated the total number of deaths for different pre-existing health conditions categories and calculated a conditional CFR based upon the number of comorbidities before SARS-CoV-2 infection morbidities before SARS-CoV-2 infection.

In the paper [13] was proved that High IL-6 level, C-reactive protein level, lactate dehydrogenase (LDH) level, ferritin level, d-dimer level, neutrophil count, and neutrophil-to-lymphocyte ratio all of them were predictors of mortality (area under the curve > 0.70), as well as low albumin level, lymphocyte count, monocyte count, and ratio of peripheral blood oxygen saturation to fraction of inspired oxygen (SpO_2/FiO_2). A multivariable mortality risk model including the SpO_2/FiO_2 ratio, neutrophil-to-lymphocyte ratio, LDH level, IL-6 level, and age was developed and showed high accuracy for the prediction of fatal outcome (area under the curve 0.94). The optimal cutoff reliably classified patients (including patients without initial respiratory distress) as survivors and nonsurvivors with a sensitivity of 0.88 and a specificity of 0.89.

As you can see there are not clearly defined factors that will affect mortality rate. There are no strict rules or decision trees for predicting patients' death.

Therefore, there is a great need to conduct research that will help the doctors predict the severity of the disease and its mortality.

The present studies and analysis unlock a way in the direction of attribute correlation, estimation of survival days, and the prediction of death probability. The findings of the present review clearly indicate that machine learning algorithms have strong capabilities of prediction and classification in relation to COVID-19 as well.

The aim of the study is the determination of the prognostic factors of fatal SARS-CoV-2-associated pneumonia and establishing a functional relationship between them and the mortality of the patient.

The main contribution of this article can be summarized as follows:

• based on the medical data of real patients of the hospital admitted with COVID-19, a heterogeneous data set was created, which became the basis for finding the relationship between the mortality rate of the patient;

• the method of validation, transformation and purification of the medical data set in preliminary preparation for the analysis was developed;

• an analysis to determine the impact of medical factors on mortality was conducted and a final set of data for the construction of classification models was formed;

• the train dataset for experimental modeling was created;

• the effectiveness of ten existing machine learning algorithms for solving the problem of determining the level of patient mortality was evaluated and a decision tree was constructed;

• a stacking model to predict mortality, which has prevented overfitting was developed and a significant increase in the accuracy of its operation and in comparison, with some existing machine learning algorithms was shown.

The resulting functional dependence can be implemented in expert systems that will allow the average physician to predict the degree of mortality of the patient, and therefore apply the necessary tools of intensive care to save human lives.

METHODS

Data Collection

A retrospective analysis of the results of treatment of 121 SARS-CoV-2-associated pneumonia patients who stayed in Chernivtsi City Hospital №1 (since March 2020 – the Chernivtsi Central COVID Hospital) was performed. The inclusion criterion was moderate or severe SARS-CoV-2-associated pneumonia as well as the exclusion criterion – the death before the fifth day staying in the hospital. According to the results, two groups were formed: the first group of the 60 SARS-CoV-2 associated pneumonia patients with the fatal outcome and the second group of the 61 patients with favorable course of the SARS-CoV-2 associated pneumonia.

Every patient could be described with a huge number of parameters. As potential prognostic factors we analyzed the 77 parameters divided into 9 parts according to the working hypothesis. This task can be attributed to the machine learning classification, where it is necessary to determine patients belonging to one of the classes (will die or live) based on many different factors. The stages of machine learning in this case should include preliminary data preparation, models selection, training and analysis of results.

Preliminary data preparation

There are several steps that are due to the peculiarities of obtaining and storing data at this stage. A Python script was written to implement each step.

Removing personalized data. Fields that contain personal information and those that do not clearly affect the diagnosis are removed from analysis. In particular: patient ID number, Name of patient, phone, diagnosis, complications, CT-scans etc.

Verification of human mistakes. The feature of the available data is that they are all entered by people, and this leads to technical mistakes. So, the first procedure is to verify the data and correct them automatically and manually. To do this, a script that identified and, if possible, corrected human errors was created.

Transformation and change of field values. A significant number of fields are not suitable for digital analysis, because they contain information in text format that is not suitable for analysis. The parse function was created that transformed all data for appropriate DataAnalysis form.

Handling features with missing data

The next step is removing the records that contain a lot of missing values. The large number of features leads to the removing records that contain at least one missing value. It can lead to a significant reduction in the DataSet and makes using classification methods impossible. To resolve this problem, empty values have been filled with the default values (if possible). Next, the features with the most missed data were identified.

It was decided to eliminate these features that consist more that 40% missed data from further calculation, as their presence will make further analysis impossible. This procedure of deletion of records with missing data reduced the DataSet by 19% (from 121 to 99 records). The total number of fields was 53 input and one output field that contain 49 – digital fields, 3 categorical and one logical.

Identification of factor importances

The Pearson's consistency criterion – χ^2 and mutual information (MI) as sorting method was used to determine the importance of factors for the classification of patients [14]–[16]. The magnitude of these criteria determines the significance of the field in the classification. The results are present in Table 1.

T a b l e 1. The top 10 of the most important features for classification

Features	χ^2	Features	MI
Leukocytes 2	434	Lymphocytes 2	0.3
Band-neutrofils 2	352	Leukocytes 2	0.28
Lymphocytes 2	352	Band-neutrofils 2	0.25
Hematocrit 2	250	Saturation without oxygen supply	0.23
Creatinine 2	226	The duration of the hospitalization	0.20
Saturation without oxygen supply	22	Hematocrit 2	0.19
The duration of the hospitalization	183	Creatinine 2	0.17
C-reactive protein 2	146	Hemoglobin 1	0.15
The pulmonary insufficiency	68	Age	0.13
Gender	50	The course of the disease	0.13

As can be seen from Table 1, the first seven factors in the two methods coincide. The only difference is their importance. Therefore, the DataSet was reduced to the first seven features.

The next step was to check the presence of correlation between features. The result of correlation analysis was presented in Fig. 1.
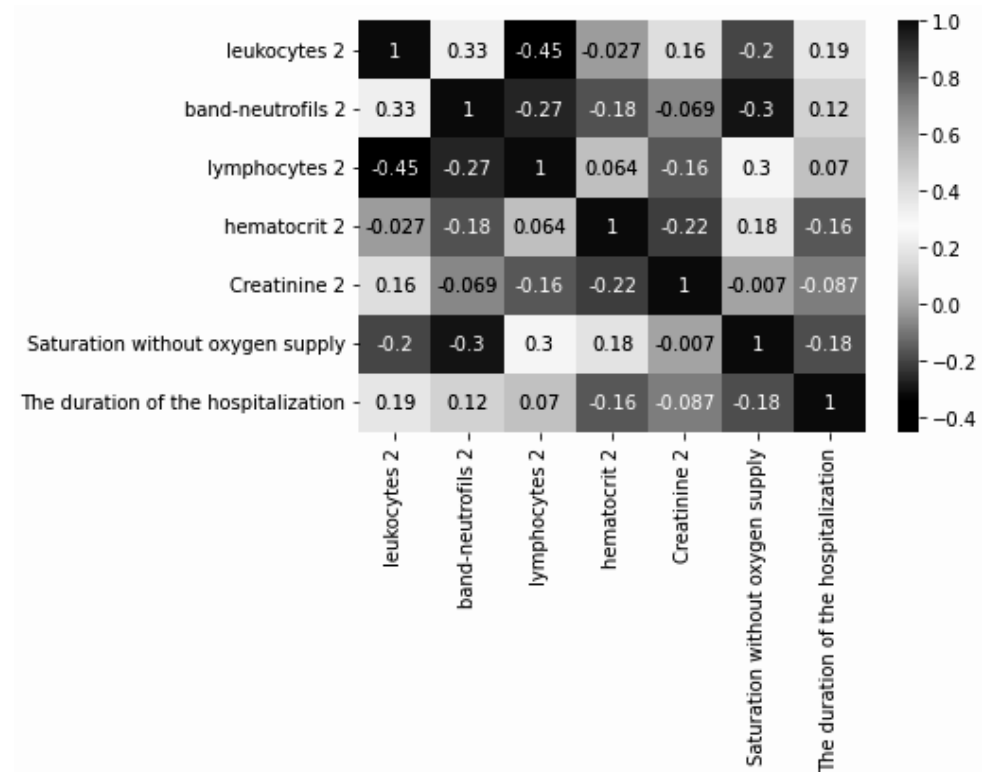

Fig. 1. Correlation matrix of input features

As can be seen from the Fig. 1, there is no correlation between the input factors. This means that you do not need to perform factor analysis and remove or convert factors.

Proposed models

This paper is aimed at building a forecast model, which will provide the highest accuracy in solving the problem on the one hand and will allow one to visualize the result in the form of a decision tree on the other hand. It is impossible to achieve this at the same time. After all, ensemble accuracy provides the highest accuracy. It is based on the use of a set of basic regressors, the results of which are summarized by the metaregressor. This will increase the accuracy compared to the use of single models that form such a model. However, it is not possible to visualize such a decision result in the form of a decision tree.

Therefore, we considered two approaches to prognosis. One is based on the decision tree; the other is ensemble.

Decision tree model. The decision tree method was used to determine the classification rules and visualize the results [17]. The main advantage of choosing this method is the ability to visualize the result of classification analysis in the form of a decision tree. However, the accuracy of this method is not the best.

The Gini coefficient was chosen as the criterion for measuring the cleavage threshold [18] – an indicator of the inequality of the distribution of some value of numbers, which takes values between 0 and 1, where 0 means absolute equality

(the value takes only one value), and 1 denotes complete inequality. The strategy used to select the split in each node is to find the best distribution.

Ensemble of classification models. The literature considers three main approaches to constructing ensemble models: boosting, bagging, and stacking.

In this work, we build a prediction model based on the stacking approach. The model assumes the presence of basic N-algorithms that will form a stacking ensemble. The meta-algorithm will weigh the results of their work. The work of the meta-algorithm will determine the impact of solving the stated task.

The data set collected by us to solve the problem of predicting the level of mortality contains many independent attributes. In addition, there are complex and nonlinear, unobvious and unexplored relationships between different features. It is evident that, in particular, many linear machine learning methods will not provide sufficient accuracy. If such algorithms are included in the general ensemble model, they will reduce the accuracy of their work. That is why we propose to perform a preliminary selection of basic algorithms that will form a stacking ensemble. It is based on initial modeling of machine learning algorithms and evaluation of their efficiency using the next four performance metrics: Accuracy, Precision, Recall and F1 Scope.

Accuracy means that the set of labels predicted for a sample must exactly match the corresponding set of labels in target.

Precision is the ratio:

$$precision = tp / (tp + fp),$$

where tp is the number of true positives and fp — the number of false positives. The precision is intuitively the ability of the classifier not to label as positive a sample that is negative.

Recall is the ratio:

$$recall = tp / (tp + fn).$$

The recall is intuitively the ability of the classifier to find all the positive samples.

F1 Scope is the harmonic mean of precision and recall, where an F1 score reaches its best value at 1 and worst score at 0. The relative contribution of precision and recall to the F1 score are equal. The formula for the F1 score is:

$$F1 = 2 * (precision * recall) / (precision + recall).$$

The Precision of classifier is the fraction of samples in the DataSet it labeled, for example, as death is really death. Its Recall is the percentage of all death samples in the dataset that it correctly labeled as death. The F1 score is the harmonic mean of precision and recall.

RESULTS AND DISCUSSIONS

Performance evaluation of the investigated decision tree model

The DataSet was splitted into train and test in the proportion of 70/30 to fit and determine the accuracy of the algorithm. The resulting decision tree is presented in Fig. 2.

Fig. 2. The decision tree determines whether the patient will die (False) or remain alive (True). Gini - distribution inequality, samples - the number of cases, value - the value of the classifier function, class - belonging to the class

Performance metrics on train DataSets consisted: accuracy = 0.90, precision = 0.89, recall = 0.88 and f1 = 0.89. We get accuracy = 0.88, precision = 0.86, recall = 0.86 and f1 = 0.86 for test DataSet. The small variance between test and training datasets indicates good fitting of this method. That is mean this model predicts unknown (new) data in the same level accuracy like know data. The high values of all metrics indicate the accuracy and adequacy of the model. It allows the doctor to personally guide the patient through this tree and quickly determine the class to which he belongs. Creating an automated decision-making program

based on trees is not a problem. The construction of the decision tree made it possible to establish the importance of features for this classifier (Table 2).

T a b l e 2. The importance of the decision tree features

Feature	Importance
Lymphocytes 2	0.62
Band-neutrofils 2	0.13
Saturation without oxygen supply	0.12
Creatinine 2	0.04
The duration of the hospitalization	0.04
Leukocytes 2	0.03
Hematocrit 2	0.02

As can be seen from the Table 2, the most important factor in the decision tree is the number of lymphocytes a week after hospitalization (lymphocytes 2). The decreased level of the lymphocytes as the marker of the severe SARS-CoV-2-infection was described in [19], [20]. Instead, our study proves the importance of this parameter as the risk marker of the fatal outcome. The further depression of the lymphocytes a week after the beginning of the intensive treatment of the SARS-CoV-2-patient points to the exhaustion of the immune defense and increases the probability of the fatal outcome.

The next important factor is the good-known indicator of the activity of the inflammatory process – the amount of the band-neutrophils [21] measured on the 7th day of the beginning of the intensive care of the SARS-CoV-2-patient. The prognostic non-favorable marker was the combination of the increasing amount of the band-neutrofils and the decreasing amount of the lymphocytes. The SARS-CoV-2-pneumonia patient's chances to survive are reduced in case of the severe activation of the inflammatory process with depression of the specific immune response.

The third important factor in the decision tree is the blood saturation without oxygen supply at the moment of the hospitalization. The low level of the blood saturation indirectly reflects the severity of the patient's condition and lungs affection, points to the exhaustion of the defensive and compensatory possibilities of the organism, the cardio-circulatory decompensation, severe tissue hypoxia [22]. The value of this indicator as a predictor of an unfavorable prognosis of the disease turned out to be quite logical.

Here are some examples of using the decision tree. The patient 1 was admitted to the hospital with blood saturation 85%, the amount of the leukocytes — 34,9 G/l, band-neutrofils — 24, lymphocytes — 3, hematocrit — 43,1, kreatinin-142 were revealed in his blood analysis in a week. Let's take the patient through the decision tree: lymphocytes ≤ 9.5 (yes) $\geq$ saturation without oxygen supply ≤ 92.5 (yes) $\rightarrow$ leukocytes ≤ 7.05 (no) $\rightarrow$ lymphocytes ≤ 7.75 (yes) $\rightarrow$ band-neutrofils ≤ 7.5 (no) $\rightarrow$ Class False, it means the prognosis is non-favorable. Indeed, on the 12th day after admission, the patient's death was fixed.

The patient 12 was admitted to the hospital with blood saturation 91%, the amount of the leukocytes — 6,1 G/l, band-neutrofils — 2, lymphocytes — 9, hematocrit — 46, kreatinin- 117 were revealed in his blood analysis in a week. Let's take the patient through the decision tree: lymphocytes ≤ 9.5 (yes) $\rightarrow$ satu-

ration without oxygen supply ≤ 92.5 (yes) → leukocytes ≤ 7.05 (yes) → saturation without oxygen supply ≤ 79.5 (no) → Class True, that shows on the favorable prognosis. And this patient was discharged from the hospital on the 13th day of the treatment.

Performance evaluation of the investigated ML-ensemble

It was decided to increase the train DataSet and use an ensemble of classification models to improve the quality of fitting and eliminate overfitting. For this purpose, all available records were used as a training set. An additional 83 patients were studied to obtain a test DataSet. New data were obtained in the same hospital department that is why the distribution of the test DataSet was the same.

The choice of classifiers for the ensemble was based on the analysis of the accuracy of each of them. The availability of overfitting on the train DataSet was also assessed. An experimental comparison of the efficiency of ten existing machine learning methods using the four performance metrics on train and test DataSets was carried out (Table 3 and 4).

T a b l e   3 . The results of prediction based on performance criteria using all the studied machine learning algorithms (Train Data Set)

Machine learning method	Performance metric			
	Accuracy	Precision	Recall	F1 Scope
Logistic regression (CR)	0.89	0.90	0.84	0.87
Decision tree (DT)	0.89	0.87	0.87	0.87
Quadratic discriminant analysis (QDA)	0.84	0.94	0.68	0.79
Naive Bayesian classifier (NB)	0.84	0.91	0.70	0.79
Random forest classifier (RF)	0.95	0.93	0.95	0.94
Adaptive Boosting classifier (AB)	1.00	1.00	1.00	1.00
Support Vector Classification (SVC)	0.89	0.95	0.80	0.86
Stochastic Gradient Descent (SGD)	0.75	0.64	0.98	0.77
Neural Network (NN)	0.98	0.97	0.97	0.97
Gradient Boosting (GB)	1.00	1.00	1.00	1.00

T a b l e   4 . The results of prediction based on performance criteria using all the studied machine learning algorithms (Test Data Set)

Machine learning method	Performance metric			
	Accuracy	Precision	Recall	F1 Scope
Logistic regression (LR) [23]	0.78	0.74	0.74	0.74
Decision tree (DT)	0.86	0.85	0.84	0.85
Quadratic discriminant analysis (QDA)[24]	0.75	0.77	0.57	0.66
Naive Bayesian classifier (NB) [25]	0.72	0.73	0.54	0.62
Random forest classifier (RF) [26]	0.62	0.56	0.57	0.56
Adaptive Boosting classifier (ABC) [27]	0.66	0.59	0.69	0.63
Support Vector Classification (SVC) [28]	0.77	0.79	0.63	0.70
Stochastic Gradient Descent (SGD) [29]	0.48	0.44	0.91	0.60
Neural Network (NN) [30–32]	0.77	0.70	0.80	0.75
Gradient Boosting (GB) [33, 34]	0.63	0.55	0.68	0.61

As you can see in Table 3, Decision Tree, AdaBoost and Gradient Boost had problems with overfitting. They have 100% accuracy on train DataSet and very low on test DataSet. Therefore, we exclude them from future analysis. All other seven classifiers had similar accuracy. Therefore, for improving accuracy we combined them into ensemble. A joint solution to these methods was found by the Voting Classifier [35]. The basic idea of a voting classifier is to combine conceptually different machine learning classifiers and use the majority of votes (hard voting) or average predicted probabilities (soft voting) to predict class labels. In our case, "hard" voting was used i.e., the choice of class was determined by the majority of "votes" of the classifiers. Results of accuracy of this ensemble are present in Table 5.

T a b l e 5 . The results of prediction based on performance criteria using ensemble of machine learning algorithms

Voting Classifier	Performance metric			
	Accuracy	Precision	Recall	F1 Scope
Train Data Set	0.94	0.95	0.91	0.93
Test Data Set	0.91	0.88	0.88	0.88

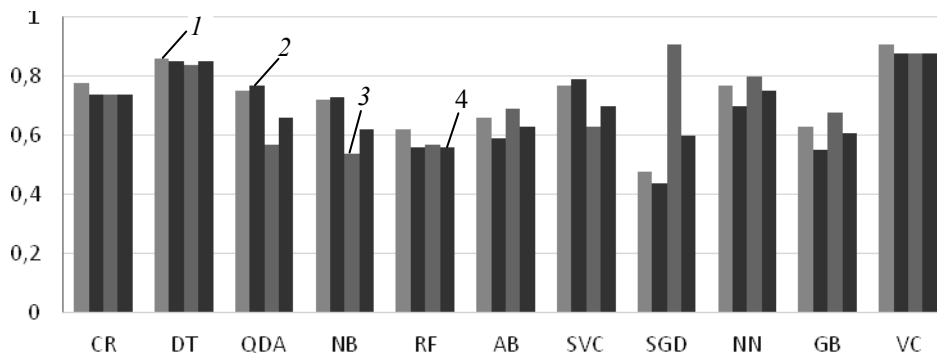For comparison we presented results on one plot (Fig. 3).

Fig. 3. Comparison of performance metrics of investigated classifiers and their ensemble: *1* — Accuracy, *2* —Precision, *3* — Recall, *4* — F1Scope

As you can see from the plot, ensemble has the biggest performance. You can also see that Recall for SGD is bigger than for ensemble. But other their performance metrics are smaller. Ensemble is stable in joint decision because Precision and Recall have the same big value. Thus, using an ensemble of ML models made it possible to avoid overfitting and increase the accuracy and stability of the forecast. The forecast error (bias) on the train DataSet is 6% and the variance of the test DataSet from the training set is 3%. So, we can conclude that to reduce the variance (reduce the error of the test DataSet) it is enough to simply increase the train DataSet. This will lead to a slight decrease in the accuracy of bias of the train DataSet and a increase in the accuracy of the test DataSet.

Further increase in the accuracy of the two indicators is possible provided the simultaneous growth of the train DataSet and the inclusion in the calculation of new factors, or the complexity of classification models, such as joining the ensemble of classifiers based on neural networks.

CONCLUSIONS

The only one marker of the non-favorable outcome of the SARS-CoV-2-associated pneumonia presented on the day of admission of the patient was the blood saturation less than 92.5%. This is the first and the basic indicator checked in every patient and doctors determine the necessity of the oxygen supply based on this parameter. In contrast to the severity of the general condition, diabetes mellitus, the duration of the disease does not increase the probability of the lethal outcome. The severity of the lung's affection based on the results of CT- or ultrasound examination don't influence the chances to die because of SARS-CoV-2-pneumonia.

But after a week of intensive treatment, we could reveal the informative markers of the lethal outcome. They are the amount of the lymphocytes and band-neutrophils in peripheral blood. The increasing of the activity of the inflammatory process reflected in the increase amount of the band-neutrophils and leukocytes as well as the decreasing of the lymphocyte points to the exhaustion of the specific immune response, the loss of the immunological control of the inflammation and to the high probability of the lethal outcome.

Using the good-known parameters that are routinely used daily in clinical practice, an accessible and understandable decision tree will allow the physician to determine the prognosis in a few minutes and, accordingly, to understand the need for treatment adjustment, transfer of the patient to the emergency department.

Creating a data collection system with further training of classifiers will dynamically increase the accuracy of the forecast and automate the decision-making process by the doctor.

DECLARATIONS

Ethics approval and consent to participate
Not applicable.
Consent for publication
Not applicable.
Availability of data and materials
Data were obtained from the medical histories of patients who were hospitalized at the Central Hospital in Chernivtsi, Ukraine. The data is available on the link: https://github.com/vyklyuk/COVID_Chernivtsi
Competing interests
The authors declare that they have no competing interests.
Funding
This research received no external funding.
Authors' contributions
Conceptualization, software, investigation, writing — original draft preparation Yaroslav Vyklyuk and Denys Nevinskyi; methodology Svitlana Levytska; software, validation, writing–review and editing Kateryna Hazdiuk; formal analysis, funding acquisition Miroslav Škoda; resources, data curation Stanislav Andrushko and Maryna Palii. All authors have read and agreed to the published version of the manuscript.

REFERENCES

1. *Worldometer COVID-19 Coronavirus Pandemic. 2020.* Accessed on: November 28, 2021. [Online]. Available: https://www.worldometers.info/coronavirus/

2. S. Priya, M. Selva Meena, J. Sangumani, P. Rathinam, C. Brinda Priyadharshini, and V. Vijay Anand, "Factors influencing the outcome of COVID-19 patients admitted in a tertiary care hospital, Madurai. -a cross-sectional study," *Clin Epidemiol Glob Health,* 2021. doi: 10.1016/j.cegh.2021.100705.

3. Annemarie Jutel, "Classification, Disease, and Diagnosis," *Perspectives in Biology and Medicine, Project MUSE*, vol. 54 no. 2, pp. 189–205, 2011. doi: 10.1353/pbm.2011.0015.

4. Aiping Lu, Miao Jiang, Chi Zhang, and Kelvin Chan, "An integrative approach of linking tradi-tional Chinese medicine pattern classification and biomedicine diagnosis," *Journal of Ethnopharmacology*, vol. 141, issue 2, pp. 549–556, 2012. Available: https://doi.org/10.1016/j.jep.2011.08.045

5. O.S. Albahri et al., "Systematic review of artificial intelligence techniques in the detection and classification of COVID-19 medical images in terms of evaluation and benchmarking: Taxonomy analysis, challenges, future solutions and methodological aspects," *Journal of Infection and Public Health*, vol. 13, issue 10, pp. 1381–1396, 2020. Available: https://doi.org/10.1016/j.jiph.2020.06.028

6. Gonçalo Marques, Deevyankar Agarwal, and Isabel de la Torre Díez, "Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network," *Applied Soft Computing*, 2020, vol. 96. Available: https://doi.org/10.1016/j.asoc.2020.106691

7. X. Wang et al., "A Weakly-Supervised Framework for COVID-19 Classification and Lesion Localization from Chest CT," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2615–2625, 2020. doi: 10.1109/TMI.2020.2995965.

8. M.E.H. Chowdhury et al., "An Early Warning Tool for Predicting Mortality Risk of COVID-19 Patients Using Machine Learning," *Cogn. Comput.*, 2021. Available: https://doi.org/10.1007/s12559-020-09812-7

9. Li Tan et al., "Validation of Predictors of Disease Severity and Outcomes in COVID-19 Patients: A Descriptive and Retrospective Study," *Med*, vol. 1, issue 1, pp. 128–138, 2020. Available: https://doi.org/10.1016/j.medj.2020.05.002

10. Ashutosh Kumar Dubey, Sushil Narang, Abhishek Kumar, Sasubilli Satya Murthy, and Vicente García-Díaz, "Performance Estimation of Machine Learning Algorithms in the Factor Analysis of COVID-19 Dataset," *Computers, Materials, & Continua*, 66(2), pp. 1921–1936, 2021.

11. Danying Liao et al., "Haematological characteristics and risk factors in the classification and prognosis evaluation of COVID-19: a retrospective cohort study," *The Lancet Haematology*, vol. 7, issue 9, pp. e671–e678, 2020. Available: https://doi.org/10.1016/S2352-3026(20)30217-9

12. M. Aguiar and N. Stollenwerk, "Condition-specific mortality risk can explain differences in COVID-19 case fatality ratios around the globe," *Public Health*, vol. 188, pp. 18–20, 2020.

13. Rocio Laguna-Goya et al., "IL-6–based mortality risk model for hospitalized patients with COVID-19," *Journal of Allergy and Clinical Immunology*, vol. 146, issue 4, pp. 799–807, 2020. Available: https://doi.org/10.1016/j.jaci.2020.07.009

14. R. Rana and R. Singhal, "Chi-square test and its application in hypothesis testing," *J. Pract. Cardiovasc. Sci.*, 1, pp. 69–71, 2015. doi: 10.4103/2395-5414.157577.

15. B.C. Ross, "Mutual Information between Discrete and Continuous Data Sets," *PLoS ONE*, 9(2), 2014. Available: https://doi.org/10.1371/journal.pone.0087357

16. E. Archer, I.M. Park, and J. Pillow, "Bayesian and Quasi-Bayesian Estimators for Mutual Information from Discrete Data," *Entropy*, 15 (12), pp. 1738–1755, 2013. doi: 10.3390/e15051738.

17. S.R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology", *Systems Man and Cybernetics IEEE Transactions*, vol. 21, no. 3, pp. 660–674, 1991.

18. Laura Elena Raileanu and Kilian Stoffel, "Theoretical Comparison between the Gini Index and Information Gain Criteria," *Annals of Mathematics and Artificial Intelligence*, vol. 41, pp. 77–93, 2004. doi: 10.1023/B:AMAI.0000018580.96245.c6.

19. J. Wagner, A. DuPont, S. Larson, B. Cash, and A. Farooq, "Absolute lymphocyte count is a prognostic marker in Covid-19: A retrospective cohort review," *Int. J. Lab. Hematol.*, vol. 42(6), pp. 761–765, 2020. doi: 10.1111/ijlh.13288.

20. A. Mazzoni, L. Salvati, L. Maggi, F. Annunziato, and L. Cosmi, "Hallmarks of immune response in COVID-19: Exploring dysregulation and exhaustion," *Semin. Immunol.*, 2021. doi: 10.1016/j.smim.2021.101508.

21. J. Wang, M. Jiang, X. Chen, and L.J. Montaner, "Cytokine storm and leukocyte changes in mild versus severe SARS-CoV-2 infection: Review of 3939 COVID-19 patients in China and emerging pathogenesis and therapy concepts," *J. Leukoc. Biol.*, 108(1), pp. 17–41, 2020. doi: 10.1002/JLB.3COVR0520-272R.

22. D. Böning, W.M. Kuebler, and W. Bloch, "The oxygen dissociation curve of blood in COVID-19," *Am. J. Physiol. Lung. Cell. Mol. Physiol.*, vol. 321(2), L349–L357, 2021. doi: 10.1152/ajplung.00079.2021.

23. J. Tolles and W.J. Meurer, "Logistic Regression: Relating Patient Characteristics to Outcomes," *JAMA*, vol. 316(5), pp. 533–534, 2016. doi: 10.1001/jama.2016.7653.

24. Alaa Tharwat, "Linear vs. quadratic discriminant analysis classifier: a tutorial," *International Journal of Applied Pattern Recognition*, vol. 3.2, pp. 145–180, 2016.

25. P. Domingos and M. Pazzani, "On the optimality of the simple Bayes-ian classifier under zero-one loss," *Machine Learning*, vol. 29, pp. 103–137, 1997.

26. Leo Breiman, "Random Forests," *Machine Learning*, 45 (1), pp. 5–32, 2001. doi: 10.1023/A:1010933404324.

27. Zhao Yan, Xing Chen, and Jun Yin, "Adaptive boosting-based computational model for predicting potential miRNA-disease associations," *Bioinformatics*, vol. 35.22, pp. 4730–4738, 2019.

28. Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin, *A Practical Guide to Support Vector Classification*. Department of Computer Science, National Taiwan University (Hrsg.), 2003.

29. L. Bottou, "Large-scale machine learning with stochastic gradient descent," *Proceedings of COMPSTAT*, Physica-Verlag HD, 2010, pp. 177–186.

30. Shaohua Wan et al., "Deep multi-layer perceptron classifier for behavior analysis to estimate parkinson's disease severity using smartphones," *IEEE Access*, 6, pp. 36825–36833, 2018.

31. D.C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Mathematical Programming*, vol. 45, pp. 503–528, 1989. Available: https://doi.org/10.1007/BF01589116

32. Ph. Moritz, N. Robert, and M. Jordan, "A linearly-convergent stochastic L-BFGS algorithm," *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics, PMLR,* 51, pp. 249–258, 2016.

33. S. Madeh Piryonesi and Tamer El-Diraby, "Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index," *Journal of Infrastructure Systems*, 26 (1): 04019036, 2020. doi: 10.1061/(ASCE)IS.1943-555X.0000512.

34. T. Hastie, R. Tibshirani, and J.H. Friedman, "Boosting and Additive Trees," *The Elements of Statistical Learning* (2nd ed.). New York: Springer, 2009, pp. 337–384.

35. Onan Aytuğ, Serdar Korukoğlu, and Hasan Bulut, "A multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification," *Expert Systems with Applications*, vol. 62, pp. 1–16, 2016.

INFORMATIONON THE ARTICLE

Yaroslav I. Vyklyuk, ORCID: 0000-0003-4766-4659, Lviv Polytechnic National University, Ukraine, e-mail: vyklyuk@ukr.net

Svitlana A. Levytska, ORCID: 0000-0001-6616-3572, Bukovinian State Medical University, Ukraine, e-mail: levitska.svitlana@bsmu.edu.ua

Denys V. Nevinskyi, ORCID: 0000-0002-0962-072X, Lviv Polytechnic National University, Ukraine, e-mail: nevinskiy90@gmail.com

Kateryna P. Hazdiuk, ORCID: 0000-0002-7568-4422, Yuriy Fedkovych Chernivtsi National University, Ukraine, e-mail: kateryna.gazdyik@gmail.com

Miroslav Škoda, ORCID: 0000-0001-6658-2742, DTI University, Slovakia, e-mail: skoda@dti.sk

Stanislav D. Andrushko, Chernivtsi central hospital, Ukraine, e-mail: stanislav.andrushko14@gmail.com

Maryna A. Palii, Chernivtsi central hospital, Ukraine, e-mail: marinapaljj90@gmail.com

МОДЕЛІ РИЗИКУ СМЕРТНОСТІ НА ОСНОВІ ДЕРЕВА РІШЕНЬ І АНСАБЛЮ ДЛЯ ГОСПІТАЛІЗОВАНИХ ПАЦІЄНТІВ ІЗ COVID-19 / Я.І. Виклюк, С.А. Левицька, Д.В. Невінський, К.П. Газдюк, М. Шкода, С.Д. Андрушко, М.А. Палій

Анотація. Присвячено вивченню пневмонії, асоційованої із SARS-CoV-2 та дослідженню основних показників, що призводять до смертності хворих. Використовуючи добре відомі параметри, які регулярно застосовуються в клінічній практиці, отримано абсолютно нові функціональні залежності на основі доступного та зрозумілого дерева рішень і моделей класифікаторів ML, що дозволить лікарю визначити прогноз за кілька хвилин і, відповідно, зрозуміти необхідність коригування лікування, переведення хворого до відділення невідкладної допомоги. Точність отриманого ансамблю моделей, підібраних за реальними даними пацієнтів лікарні, становила 0,88–0,91 для різних показників. Створення системи збирання даних з подальшим навчанням класифікаторів дасть змогу динамічно підвищити точність прогнозу та автоматизувати процес прийняття рішення лікарем.

Ключові слова: COVID-19, система прийняття рішень, дерево рішень, ML-ансамбль, ансамбль класифікаційних моделей.